A Road Edge Detection Approach for Marked and Unmarked Lanes Based on Video and Radar

Florian Janda, Sebastian Pangerl, Andreas Schindler
Institute for Software Systems in Technical Applications (FORWISS)
University of Passau
D-94032 Passau, Germany
{jandaf, pangerl, schindler}@forwiss.uni-passau.de

Abstract—In this paper we present a reliable and real-time perception approach for detecting road boundaries by fusing video and radar. The road boundary is defined as the change from road surface to non-road area. We show the integration of a multi-lane recognition system into the detection algorithm, which makes the approach independent of the number of lanes and the visibility of lane markings. The performance of the system is evaluated by means of manually labeled reference data. Information on the road geometry like the curvature and the relative position of the ego vehicle is crucial for ADAS. The road boundary detection acts as main component for postponed functions like a run-off-road prevention, which keeps a vehicle on the drivable area. This work is part of the European project interactIVe, which addresses new technologies and approaches to increase vehicle safety through an integrated platform.

I. INTRODUCTION

Safety technologies in terms of advanced driver assistance systems (ADAS) are starting to contribute to casualty reduction and hold potentially large future promise. In recent years, research was very active in order to increase traffic safety. A special type of accidents are single vehicle crashes in which no other road user is involved. These accidents include run-off-road collisions, rollover crashes and collisions with a solid obstacle like a tree when traversing the roadway unintentionally. An improvement of road infrastructure and the installation of safety facilities like guardrails contribute to decrease the fatalities by passive safety. Even if the statistical data kept by the European union identifies a decrease of 36% of single vehicle accidents in the last ten years, single vehicle crashes are still responsible for one third of all traffic accident fatalities in countries of the EU [1]. Thereby, country roads and non-urban areas account for the most affected roads. Another domain that profits from reliable road boundary detection is autonomous driving. Such systems require an accurate environment perception of the road geometry and the relative pose of the vehicle on the road. Escape path strategies in cases of unavoidable crashes are also dependent on information about the drivable road area. In this work we present a real-time capable perception approach for road boundary detection, where the road boundary, or the road edge, is defined as the change from road surface to non-road area. We interpret a structural separation like guardrails or other barriers flanking the road side as a road boundary, as well. We use a particle filter [2] to identify hypothesis for the road edge by analyzing image and radar features. Integrating a multi-lane detection system makes the algorithm unaffected by the road type. Finally, we present some examples and evaluate the road edge detection regarding to different scenarios.

In general, the road boundary estimation is a challenging task. It has to deal with different types of roads as highway, country, rural, tunnel, etc., where lane markings are visible, scarce or missing at all. Additionally, illumination changes like shadows or irregularities on the road surface make a separation from road to non-road area very difficult. A lot of approaches are given in literature that deal either with multi-lane recognition on well structured roads [3] or focus on country roads with vanishing markings [4]. Detecting the road boundary in case of arbitrary combinations of lane markings and unmarked road sides is a challenging task. We face this problem by combining the techniques of a multi-lane detection and road segmentation in order to determine the road boundary on asphaltic roads. The proposed method is not restricted to a specific road type and needs only little supplementary knowledge from digital maps at unmarked roads for an initial guess of the road width.

A popular technique to separate the homogeneous road surface from the non-road area is texture analysis (cf. [5], [6]). The authors in [4] transform texture information into gray values and apply a N-level fit with geometric constraints. The approach presented in [7] takes additional features like color and edge strength information in order to estimate the road parameters. Both approaches are designed for country roads and might not return the road boundary if markings are present. Although color information seems to be a promising feature for road segmentation (cf. [8], [9]), we use a monochrome camera only. Instead of feature based approaches like [10], that require a lot of training data for road segmentation, we rely on the estimation of a geometric road model by online calculated image features. Additionally, due to the fact that installed barriers along the road side are promising indicators of the road endings [11], we will present how to integrate a distance sensor like a radar into the perception process. In the following we define the used edge model and present our algorithm based on particle filtering. Then we show how subparts of the approach can be evaluated. Finally, we demonstrate the performance of our algorithm on some real road scenarios.

II. ROAD EDGE DETECTION

We present a probabilistic approach to estimate the parameters of a geometric model that represents the road edge. It is based on the concept that was introduced in [7]. The main advantage of our work is a masking strategy for image regions,
that have been identified by a preceding multi-lane recognition system and a sophisticated texture analysis. Hence the stability and quality of the detection of the road boundary is improved. Additionally, we introduce new image and radar features to the fusion process. Before scooping on the estimation method, we define the considered road representation.

A. Road Edge Model

Our approach describes the road edge with a geometric model in the street coordinate system, where the x-axis points in the longitudinal direction of the vehicle and the y-axis to the left. The usage of a geometric model has proven to be appropriate in many lane recognition systems [12]. As shown in [13], the Taylor approximation of a clothoid leads to a third order polynomial. We obtain a polynomial

\[ \hat{e}_l : [0, \infty) \rightarrow \mathbb{R}, x \mapsto \frac{c_1}{6} x^3 + \frac{c_0}{2} x^2 + \varphi x + o + \frac{r}{2} \]

for the left road edge, with the yaw-angle \( \varphi \) between the longitudinal vehicle axis and the tangent at the start point, the curvature \( c_0 \), the curvature rate \( c_1 \), the width of the road \( r \) and \( o \) representing the oriented distance of the center of the road to the mid of the vehicle. For the right road edge \( \hat{e}_r \) the term \( \frac{r}{2} \) is replaced by \( -\frac{r}{2} \).

B. Algorithmic description

The parameters of the street model form the considered state space of the estimation problem. As usual in a particle filter approach, the state probability density is approximated discretely by a finite set of particles with corresponding weights. Each particle represents a state hypothesis, while the weight corresponds to the likelihood for observing some measurement given this state hypothesis. The sequential Monte-Carlo method recursively updates the weights with respect to new measurements followed by a redistribution of particles according to a resampling strategy based on the updated weights. Each of the \( n \) particles of the form

\[ p_i = (\varphi^{(i)}, c_0^{(i)}, c_1^{(i)}, r^{(i)}, o^{(i)}), i \in \{0, \ldots, n-1\} \]

describes a hypothesis for an edge model state. An evaluation function \( w \) assigns a weight to each particle \( p_i \). The weight for a set of image and radar features is given by the product of the individual feature weights. Finally, the best road edge hypothesis is the mean over all particle states. Fig. 1 illustrates the algorithm overview, that mainly consists of a preprocessing block and the particle weighting functions. In the next section we explain in detail how the input images are derived from the raw image and we describe the individual particle weighting functions of each feature.

1) Texture Filtering: We use a two-dimensional DCT to identify image parts with high and low texture information. The DCT is well known as a tool for analyzing local frequency structures in a signal. Starting with the raw camera image \( I_{\text{RAW}} \), restricted to a region of interest defined by the visible road, a \( 8 \times 8 \) sized window is slid over \( I_{\text{RAW}} \) with an offset of three pixels. For each window a 2d DCT is applied. Since we are interested in strong gray value changes, which are encoded in the higher frequency part of the signal, the first coefficient of the DCT is set to zero because it represents the mean gray value. The new gray value for the center pixel of the window in \( I_{\text{RAW}} \) is the squared sum of the remaining 63 coefficients of the DCT. \( I_{\text{DCT}} \) has a third of the size of \( I_{\text{RAW}} \) due to the pixel shift of the DCT window. The parts of \( I_{\text{RAW}} \) with low texture information, like the street surface, contain low gray values in \( I_{\text{DCT}} \). In turn, the parts with significant texture changes, like the edge of the road, lead to high gray values. Examples for that characteristic are given in Fig. 2. We denote the width of \( I_{\text{DCT}} \) by \( i_w \) and the height by \( i_h \).

\[ P = \{0, \ldots, i_w - 1\} \times \{0, \ldots, i_h - 1\} \]

is the set of all pixel indices. Furthermore, in order to reduce the computational time of the subsequent image processing methods, we consider a shrinked image \( I_S \) of \( I_{\text{RAW}} \) with the size of \( I_{\text{DCT}} \).

2) Map Information: Our algorithms are designed to work with few additional information. However, the integration of a map service helps to improve the detection quality. Solely we
use information on the road type to initially estimate the width of the road after reinitializing the particle filter. By default, it is set to 6.5\textit{m} on highways and to 3\textit{m} on all other roads. If the map service is not available we try to estimate the width of the road by using the detected road markings. However, the road may be unmarked or no markings can be detected, then we assume a minimal road width of 3\textit{m}. Without a good initial guess of the width of the road it may take longer to find the road edge.

3) Road Marking Detection and Image Masking: The approaches in [4], [7] focus on unmarked roads. Applying this algorithms on marked roads may let the edge estimation get stuck to road markings as the image features cannot distinguish between road markings and road edges. We face this problem by masking image regions that are likely to contain road markings. Therefore, we adapted and used the multiple lane detection system presented in [3]. This approach works without any a priori knowledge on the road type or the number of present lanes. Subsequently, we create a binary image \( I_{\text{MASK}} \) of the same size as \( I_{\text{DCT}} \) and \( I_{S} \) where only the detected road markings and their neighborhood are encoded by ones. Masked image areas are treated as areas with “no information” and will be ignored in the feature weighting step. We call

\[
B_M = \{(u, v) \in P \mid I_{\text{MASK}}(u, v) = 1\}
\]

the set of masked pixels. The result of the road marking detection and a mask image is shown in Fig. 3.

Furthermore, the information provided by the road marking detection is used to complement the data from the map service, when only incomplete information about the number of lanes or the width of the road is available. Naturally, this is only possible if the road is marked. At first, the so called ego lane is estimated. This is the lane the vehicle is driving on. Commonly, it is bordered by two adjacent road markings. The width of the ego lane is used as lower bound for the distance of the two edges. In addition, the knowledge about the number of lanes and the width of each road can be integrated in the resampling step of the particle filter.

C. Particle Weights

The features can be divided into three groups: The first features are based on a shrinked version \( I_{S} \) of the input image \( I_{\text{RAW}} \), the second features use the texture filtered image \( I_{\text{DCT}} \) and the third features referring to radar measurements. The geometric edge models \( e_l \) and \( e_r \), parameterized by a particle \( p_i \), are sampled up to a distance of \( 30m \) in one meter steps. The sampled points are projected into the preprocessed images \( I_{\text{DCT}} \) and \( I_{S} \). By interpolating the projected sampled points at \( v \in \{0, \ldots, i_h - 1\} \) we denote the corresponding value of the left edge by \( e_l(v) \) and by \( e_r(v) \) for the right edge, respectively. Note that there might be particle states where at least one edge \( e_l(v) \) or \( e_r(v) \) is not in \([0, i_w - 1]\) for some \( v \). The latter corresponds to road hypotheses where the road edge is only partly visible or not visible at all. We denote all pixel belonging to an edge boundary by

\[
B_I(p) = \{(\text{round}(e_l(v)), v)\mid v \in \{0, \ldots, i_h - 1\}\} \cup \{(\text{round}(e_r(v)), v)\mid v \in \{0, \ldots, i_h - 1\}\}.
\]

For features based on the data of a radar, no coordinate transformation is needed as both the data and the edge model are given in street coordinates. We observed that image features based on contours of the \( I_{S} \) and \( I_{\text{DCT}} \) as well as features based on two level fitting on \( I_{\text{DCT}} \) performed best. In the following the calculation of these features will be described.

1) Edge Gradient Feature: The motivation for this feature is that contours in \( I_{S} \) or \( I_{\text{DCT}} \) are significant indicators for the road edge. Therefore, the weight of a particle should be the higher the closer the road edges \( e_l \) and \( e_r \) are to the nearest contour and the higher the gradient of that contour is. In [7] a function \( \omega \) is presented that assigns to each pixel \((u, v)\) an energy value depending on the distance to the closest contours and the absolute gradient value of that contour. Finally, the total weight for a particle \( p \) is given by

\[
w_{\text{edge}}(p) = \prod_{(u,v) \in B_I} \omega(u,v).
\]

2) Texture Feature: Another indicator considers the texture between the edges and outside the edges. We assume that the noise level or the change in texture is lower in the on-road area than in the off-road area. As mentioned, the DCT image highlights image areas with much change in texture. This behavior is shown in Fig. 2. Hence, we use \( I_{\text{DCT}} \) as base image. According to an edge hypothesis given by a particle \( p \), the image is separated in an on-road area and an off-road-area. We are interested in the mean squared gray value in the off-road area. In a first step an integral image \( I_{\text{INT}} \) based on \( I_{\text{DCT}} \) is calculated. Each pixel of the integral image contains the row-wise gray value sum of all pixels from the start of the row. Remember that we ignore those pixels that are masked in \( B_M \). Furthermore,

\[
g_{\text{sum}} = \sum_{(u,v) \in P \setminus B_M} I_{\text{DCT}}(u,v)
\]

is the sum of all gray values in the image (except those belonging to road markings)

\[
n_{\text{sum}} = \text{card}(P) - \text{card}(B_M)
\]
and represents the total number of unmasked pixels. We define a function

$$\delta(v) = \begin{cases} 1 & \text{if } e_x(v) \geq 0 \land e_y(v) < i_w \\
0 & \text{else} \end{cases}$$

which indicates that a road edge hypothesis is visible in one row of $I_{DCT}$. Now we sum up all gray values between the road edges line by line

$$g_{road} = \sum_{v \in \{0, ..., i_k \}} \delta(v) \cdot (I_{INT}(e_x(v), v) - I_{INT}(e_y(v), v)).$$

The number of pixels in all rows that count to the on-road area and represent no road markings are denoted by $n_{road}$. Finally, we obtain the average squared gray value as

$$w_{texture}(p) = \frac{(g_{sum} - g_{road})^2}{n_{sum} - n_{road}}$$

which is used as weight.

3) Radar Feature: A radar provides a set $M$ of measurements $\hat{m} = (m_x, m_y) \in \mathbb{R}^2$. We separate the measurements according to the best road edge hypothesis in measurements $M_E$ belonging to boundaries on the left side and $M_R$, respectively. Furthermore, an upper threshold $t$ limits the maximum distance of a measurement to a boundary. We have chosen $t = 10m$. Then the distance of a measurement to an edge $\hat{e}_l$ or $\hat{e}_r$ is calculated using

$$d(\hat{e}, \hat{m}) := |m_y - \hat{e}(m_x)|.$$

Each measurement is weighted by

$$\eta(\hat{e}, \hat{m}) := \begin{cases} (t - d(\hat{e}, \hat{m}))^2 & \text{if } d(\hat{e}, \hat{m}) < t \\
0 & \text{else} \end{cases}$$

The weight for a distance sensor measurement is then

$$w_{radar}(p) = \sum_{\hat{m} \in M_E} \eta(\hat{e}_l, \hat{m}) + \sum_{\hat{m} \in M_R} \eta(\hat{e}_r, \hat{m}) + 1.$$

Finally, the weight of a particle $p$ is given by the product of two edge gradient features, a texture feature and a radar feature:

$$w(p) = w_{edges}(p) \cdot w_{edge,DCT}(p) \cdot w_{texture}(p) \cdot w_{radar}(p).$$

The mean over all particles is interpreted as the best road edge hypothesis.

D. Feature Evaluation

One main profit of our approach is the incorporation of road markings. Now we face the performance of the image features with and without the compensation of road markings. The output of a single image feature is shown in Fig. 4 for a simplified road geometry. We have chosen a straight road segment as depicted in Fig. 3. Hence, the yaw angle, the curvature and the curvature rate of the edge model are assumed to be zero. The images in Fig. 4 is interpreted as follows: Each pixel corresponds to a pair of a road width $r$ and an offset value $o$. A single row in the image corresponds to a fixed road width and offset values from -5m on the left image border to 5m on the right image border. In turn, each column in the image corresponds to a fixed offset value and road width values from 2m in the top image border to 12m on the bottom image border. The gray value of the image visualizes the feature weight assigned to a road model parametrization defined by the road width and offset corresponding to the pixel coordinate (high gray values correspond to high feature weights). For instance, the gray value at the center of such an image represents the weight of a particle with offset 0m and width 7m. The gray value at the center top of the image represents the pair $o = 0m$ and $r = 2m$, respectively.

One can see that there are several combinations of width and offset, that lead to high weights. Diagonal lines appear whenever a road edge referring to a parameter pair hits a real road edge. The intersections of such diagonal lines belong to combinations of width and offset that are likely to represent a pair of real road edges.

In the first row the weight distribution without road markings compensation is shown. In contrast, in the second row the road markings compensation was enabled. The images with activated road markings compensation show less combinations with high weights, as the road marking candidates are removed. This improves significantly the probability that the detector is stuck to the real road edge, than to an edge, that is caused by a road marking. Especially the texture feature looks a lot more cleaned-up. Also the edge feature on $I_{DCT}$ shows only one explicit maximum for the given example. Further evaluation showed a significantly reduction of falsely detected road edges with an activated road markings compensation.

III. EVALUATION

Our work is part of the interactIVe perception platform, which is used by Volvo Truck, Ford and Fiat car in the project. Two test vehicles are equipped with a gray value camera with 30 frames per second and a resolution of $752 \times 480$. The camera in the third test vehicle has a resolution of $700 \times 250$ pixel. All platforms offer access to radar data. In addition the Fiat vehicle is equipped with a laser scanner, that has been used for testing purposes, as well. Our algorithms are running on different car
PCs in real-time. A trade-off between computation time and detection quality is achieved by using 200 particles.

The road edge detection was evaluated on different vehicles using the perception platform, among others a truck. The approach was faced with a lot of difficult scenarios including roads with and without markings, wide roads like highway, ramps and tunnels. We observed also a fast adaption of the road edge detection in case of short road extensions like bus stops. No map-based ground-truth data was available for the evaluation. Therefore, the road geometry including the road edges and the road markings was labeled manually in the video. We have selected reference points in each fifth frame in order to keep the effort limited. For the evaluation, these points are re-projected to the street plane and the distances to the estimated road edge are measured. Until now, we do not use any pitch compensation. We expect improved results, especially in the truck, as the driver’s cab experiences an intensive pitch.

We have analyzed the absolute and oriented mean error of the re-projected reference points to the assumed road edge. By \( \hat{R}_l \) and \( \hat{R}_r \) we denote the sets of reference points in the street coordinate system that belong to the left or the right edge and \( \hat{r} = (r_1, r_2) \) is a single reference. Then the absolute mean error for a frame is denoted by

\[
E = \frac{\sum_{r \in \hat{R}_l} |r_2 - e_l(r_1)|}{\text{card}(\hat{R}_l)} + \frac{\sum_{r \in \hat{R}_r} |r_2 - e_r(r_1)|}{\text{card}(\hat{R}_r)}.
\]

The oriented mean error is evaluated separately for each road edge:

\[
E_l = \frac{\sum_{r \in \hat{R}_l} r_2 - e_l(r_1)}{\text{card}(\hat{R}_l)}, \quad E_r = \frac{\sum_{r \in \hat{R}_r} r_2 - e_r(r_1)}{\text{card}(\hat{R}_r)}.
\]

Hence, we can give a statement whether a single estimated road edge is systematically on the wrong side of the real road edge.

Exemplarily, we describe the result for a rural road driven with a truck. The recorded scene has a length of about 70s. At the beginning the road is straight and is followed by a long curve. The following figures refer to this test scenario. The labeled reference points are illustrated in Fig. 5. Most of the time, more than two reference points are available. For evaluation only points up to a distance of 30m are considered. In Fig. 6 we compared the curvature of the road estimated by the algorithm with the curvature derived from the inertial measurement unit. Furthermore, Fig. 8 denotes the detected distances to the left and the right road edge. It turned out that the average error \( E \) is below 25cm for this test scenario and \( E_l \) and \( E_r \) are about -12cm. Fig. 7 depicts the mean absolute error. Analyzing several tracks, we can conclude that the average error for the closer road edge (often, this is the right one) is better than for the opposite one. This is caused by the fact, that the left road edge is only visible at a distance of about 20m for wide roads. In contrast, the other edge is visible starting at 5m. Further evaluation showed that the oriented errors \( E_r \) and \( E_l \) differ significantly from 0 especially in sections with guardrails. We found an explanation for this behavior: In cases where the road is flanked with guardrails, the radar feature overrules the image feature. Then, sometimes, the road edge is not the transition from asphalt to grass that was labeled but the projection of the guardrail to the street plane. Fig. 9 shows the performance of the road edge detection on different scenarios. Images a) - f) show successful road edge detections on marked and unmarked rural roads, on highways and ramps with and without guardrails. Although e) and h) look very similar, the detection fails in h). In h) the road edge is next to the outer road marking. Hence, the mask image hides the real road edge and the detection is stuck to a texture irregularity on the left lane and the ditch on the right. In e) there is enough asphalt visible between the marking and the road edge. By contrast, in g) the width of the road markings is underestimated and the marking is not masked completely. Therefore, the right border of the marking is interpreted as the best edge hypothesis.
Fig. 9. The performance of the road edge detection on different scenarios. a) - f) show correct boundary estimation. In g) the right boundary sticks to a lane marking and in h) the algorithm decides for the ditch as best edge hypothesis, whereas on the opposite lane the left boundary is caught by road irregularities.

Problems occur within construction areas or on partially renewed roads that make the texture non unique for the whole road area. Furthermore, high traffic density, that occludes the boundary for a long time degrades the quality as well.

IV. Conclusion

The concept of a robust road boundary detection based on a particle filter framework is presented in this paper. Texture analysis proved to be appropriate for the separation of road surface and non-road regions. However, the existence of visible lane markings makes the features derived from texture not suitable. For this reason we use a multi-lane recognition for masking these image areas in order to cope with the noise in texture transform produced by white markers. Motivated by the presence of stationary road barriers we incorporate a distance sensor like radar as well. The outcome is a realtime capable perception approach for road boundary detection that works on different kind of roads without any supplementary knowledge about the road infrastructure. The algorithm is integrated into a perception platform running in parallel with a lot of other perception modules and the provided information is made available for various functions. This includes the avoidance of run-off the road accidents by active steering. The information provided about the road geometry is used for reliable escape path planning in case of rear-end collisions or a forthcoming unwanted left of the driving surface.

In the future we will further improve the features and work on the road marking masking to succeed in scenarios as depicted in Fig. 9 g) and h). Particularly, the weighting strategies between different features will be examined. We will investigate in other sophisticated image features improving the performance, that can be easily evaluated on the labeled road scenarios. We also plan to integrate a ego motion estimation to cope with heavy pitching of the vehicle.

Acknowledgment

This work was also supported by the European Commission under interactIVe, a large scale integrated project part of the FP7-ICT for Safety and Energy Efficiency in Mobility. The authors would like to thank all partners within interactIVe for their cooperation and valuable contribution.

References